4-18-2007

Semantic Web Applications in Industry, Government, Health Care and Life Sciences

Amit P. Sheth
Wright State University - Main Campus, amit@sc.edu

KNO.E.SIS

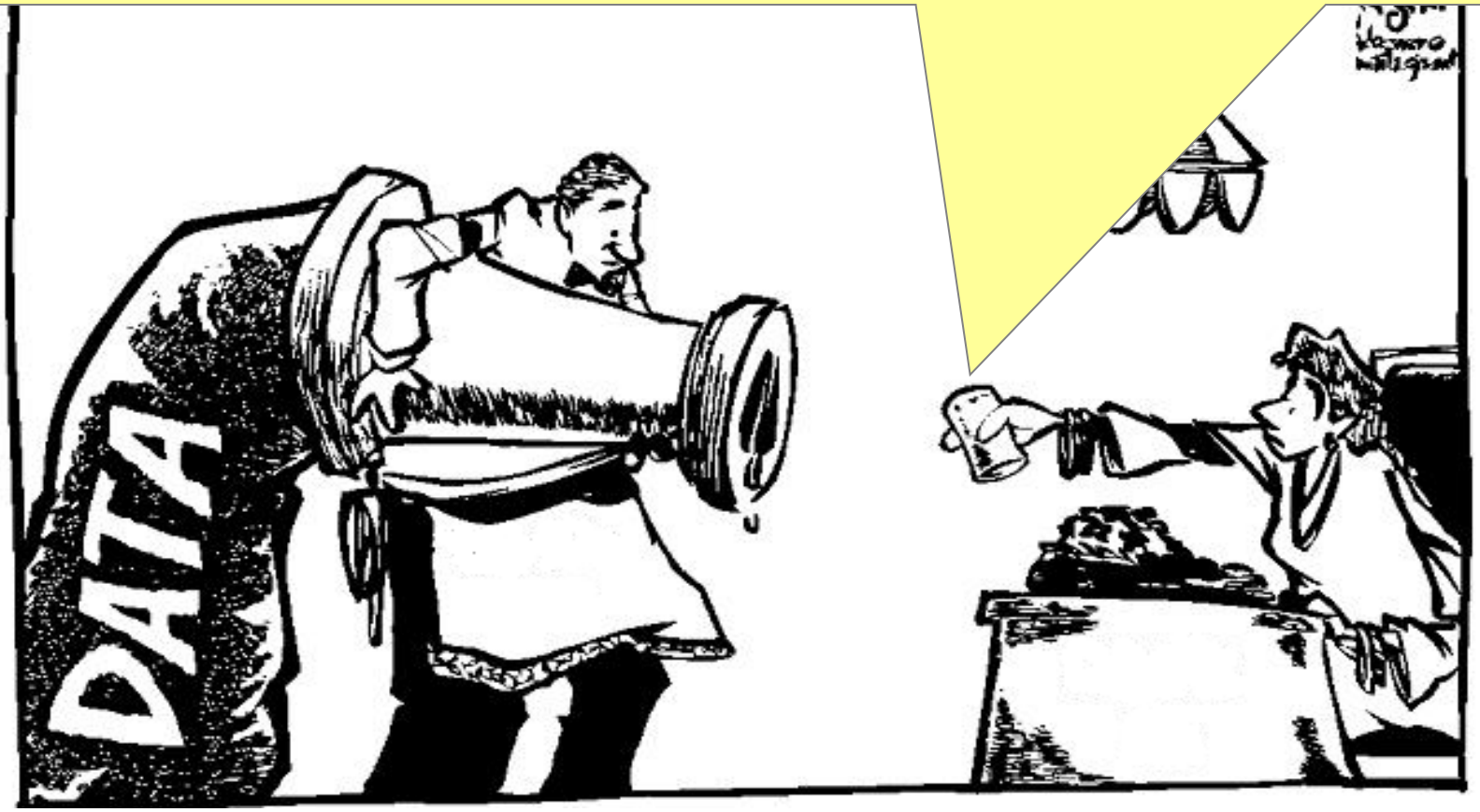
COLLECTING THE DOTS | CONNECTING THE DOTS

Semantic Web applications in Industry, Government, Health care and Life Sciences

Amit Sheth

- Death by data: amount, variety, sources
- How to exploit this data?
- Silver Bullet – SEMANTICS (approach) & Semantic Web (technology)
- Applications to show the value
  - Industry (financial services)
  - Government (intelligence)
  - Science (health care and biomedicine)

Not data (search), but integration, analysis and insight, leading to decisions and discovery



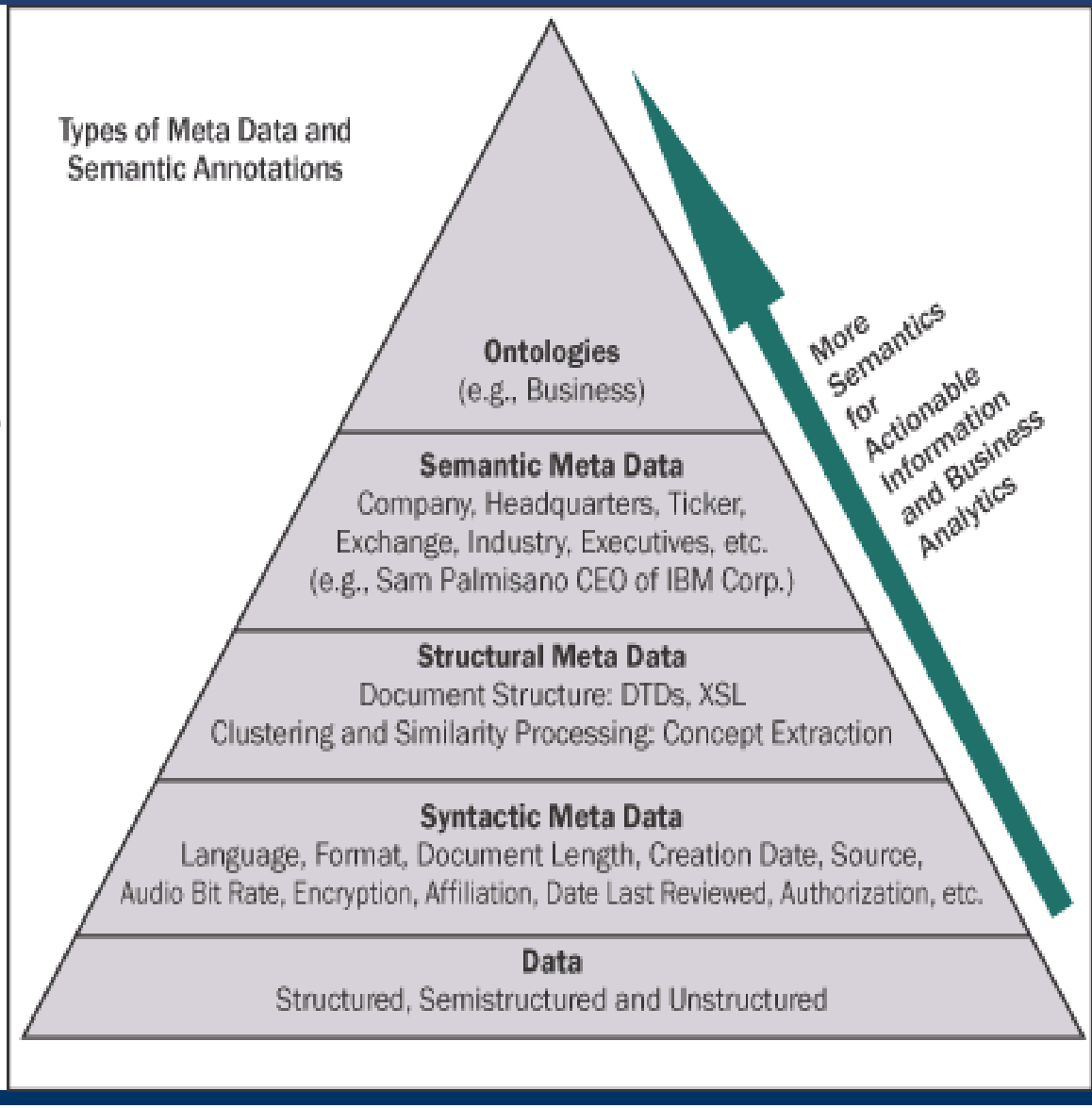
- Data captured per year = 1 exabyte (10^{18})
  (Eric Neumann, Science, 2005)
- How much is that?
  - Compare it to the estimate of the total words ever spoken by humans = 12 exabyte
- Continued high cost of interoperability and integration
- Needle in the haystack

Death by Data: Increasing Heterogeneity & Complexity

- Multiple formats: Structured, unstructured, semi-structured
- Multimodal: text, image, a/v, sensor, scientific/engineering
- Thematic, Spatial, Temporal
- Enterprise to Globally Distributed

What?

Moving from Syntax/Structure to Semantics



Semantics: **Meaning & Use of Data**

Semantic Web: **Labeling data on the Web so both humans and machines can use them more effectively**

i.e., Formal, machine processable description ⇨ more automation;

emerging standards/technologies

(RDF, OWL, Rules, ...)

Semantic Annotation - document

How?

Ontology: Agreement with Common Vocabulary & Domain Knowledge

Semantic Annotation: metadata (manual & automatic metadata extraction)

Reasoning: semantics enabled search, integration, analysis, mining, discovery

- Agreement
- Common Nomenclature
- Conceptual Model with associated knowledgebase (ground truth/facts) for an industry, market, field of science, activity
- In some domains, extensive building of open-source ontologies, in others, build as you go

Semantic Annotation – news feed

<Entity id="494805" class="DrugOntology#Organization"
Public Health Advisory FDA Announces Important Change
<Entity id="492805" class="DrugOntology#prescription_
Selective and Non-Selective Non-Steroidal Anti-Inflamma
class="DrugOntology#prescription_drug_generic">NSAI

Today, the Food and Drug Administration (FDA) is announ
<Entity id="474305" class="DrugOntology#Organization"
to voluntarily withdraw <Entity id="122805" class="
DrugOntology#prescription_drug_brandname">Bextra
</Entity> (<Entity id="10288" class="DrugOntology#pre
valdecoxib</Entity>) from the market. Pfizer has agreed to
of Bextra in the <Entity id="7852" class="Sweto#country"
further discussions with the agency.

.....

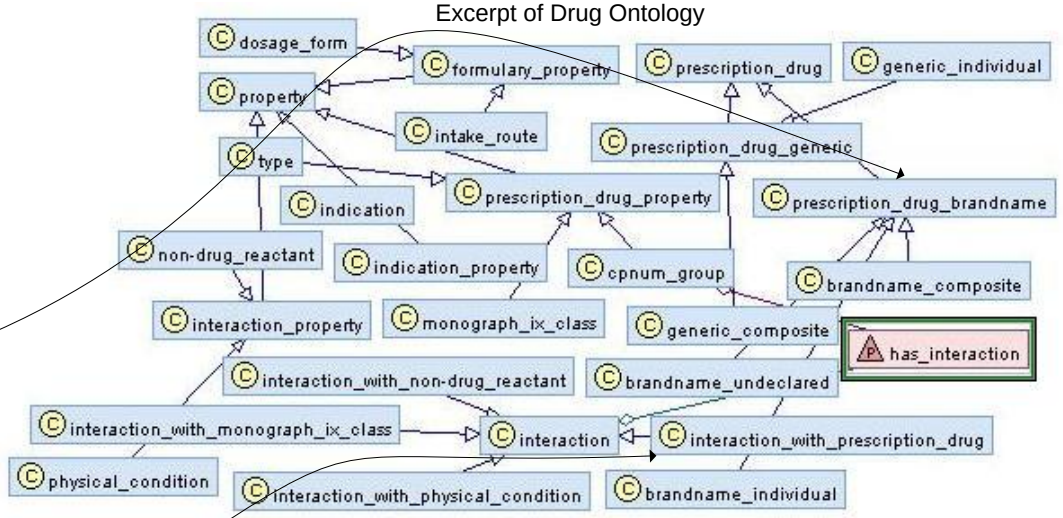
This request is based on:

-Reports of serious and potentially life-threatening skin reactions, including deaths, in patients using Bextra. The risk of these reactions in individual patients is unpredictable, occurring in patients with and without a prior history of <Entity id="14280" class="DrugOntology #interaction_with_physical_condition>sulfa allergy
</Entity>

.....

Date created: <Regexp type="date">April 7, 2005</Regexp>

Excerpt of Drug Ontology



Sample Created Metadata

<Entity id="122805"
class="DrugOntology#prescription_drug_brandname">
Bextra
<Relationship id="442134"
class="DrugOntology#has_interaction">
<Entity id="14280" class="DrugOntology
#interaction_with_physical_condition>sulfa allergy
</Entity>
</Relationship>
</Entity>

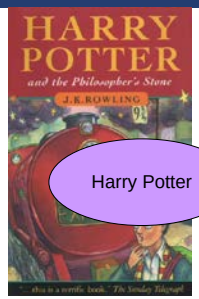
A scenario

Joe, a high school student researching the pseudo-history behind The *Da Vinci Code* . Teacher suggests him to focus on:

- The Last Supper by Leonardo da Vinci
- Et in Arcadia ego by Nicolas Poussin

It so happens that Joe's younger brother is reading *Harry Potter and the Philosopher's Stone* by J.K. Rowling. A casual conversation between the brothers leads to a discussion about Nicolas Flamel, who in the Harry Potter novel is said to have possessed the philosopher's stone that could convert base metals to gold. At this point their father, an amateur historian suggests that there might have been a real person by the name Nicholas Flamel, and the three wonder if there is any ***link*** between these seemingly unrelated stories.

Anecdotal Example

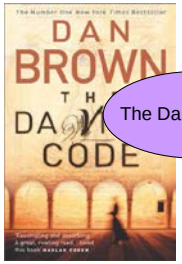


Harry Potter

UNDISCOVERED PUBLIC KNOWLEDGE

Discovering connections hidden in text

Et in Arcadia Ego



The Da Vinci code

Santa Maria delle Grazie

Priory of Sion - Wikipedia, the free encyclopedia - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://en.wikipedia.org/wiki/Priory_of_Sion

Customize Links Free Hotmail My Yahoo! Windows Marketplace Windows Media Windows Yahoo! Bookmarks Yahoo! Mail Distributed data stru... Yahoo!

Search MSD Text Search for OR Get PDB:

The Priory of Sion was supposedly led by a Grand Master or Nautonnier.

1. Ugo de Blancheford (1150-1151)
2. Bernard de Tremblay (1151-1153)
3. Guillaume de Chanaleilles (1153-1154)
4. Evrard de N...? (1154-1154)
5. Andrè de Montbard (1155-1156)
6. Bertand de Blanchefort (1156-1169)
7. Philippe de Milly (1169-1170)
8. Eudes de Saint-Amand (1170-1180)
9. Arnaud de Toroge (1181-1184)
10. Gérard de Rideford (1184-1188)
11. Jean de Gisors (1188-1220)
12. Marie de Saint-Clair (1220-1266)
13. Guillaume de Gisors (1266-1307)
14. Edouard de Bar (1307-1336)
15. Jeanne de Bar (1336-1351)
16. Jean de Saint-Clair (1351-1366)
17. Blanche d'Evreux (1366-1398)
18. Nicolas Flamel (1398-1418)
19. Rene d'Anjou (1418-1480)
20. Iolande de Bar (1480-1483)
21. Sandro Filipepi AKA Botticelli (1483-1510)
22. Leonardo da Vinci (1510-1519)
23. Charles III (Duke of Bourbon-Montpensier) (1519-1527)
24. Ferdinand de Gonzague (1527-1556)
25. Michel de Notre-Dame AKA Nostradamus(1556-1566)
26. Duc de Longueville & Nicolas Froumenteau (1566-1575)
27. Louis de Nevers (1575-1595)
28. Robert Fludd (1595-1637)
29. Johann Valentin Andrea (1637-1654)
30. Robert Boyle (1654-1691)
31. Isaac Newton (1691-1727)
32. Charles Radclyffe (1727-1746)
33. Charles de Lorraine (1746-1780)
34. Maximillian de Lorraine (1780-1801)
35. Charles Nodier (1801-1844)
36. Victor Hugo (1844-1885)
37. Claude Debussy (1885-1918)
38. Jean Cocteau (1918-1963)
39. Pierre Plantard (1963-1981)

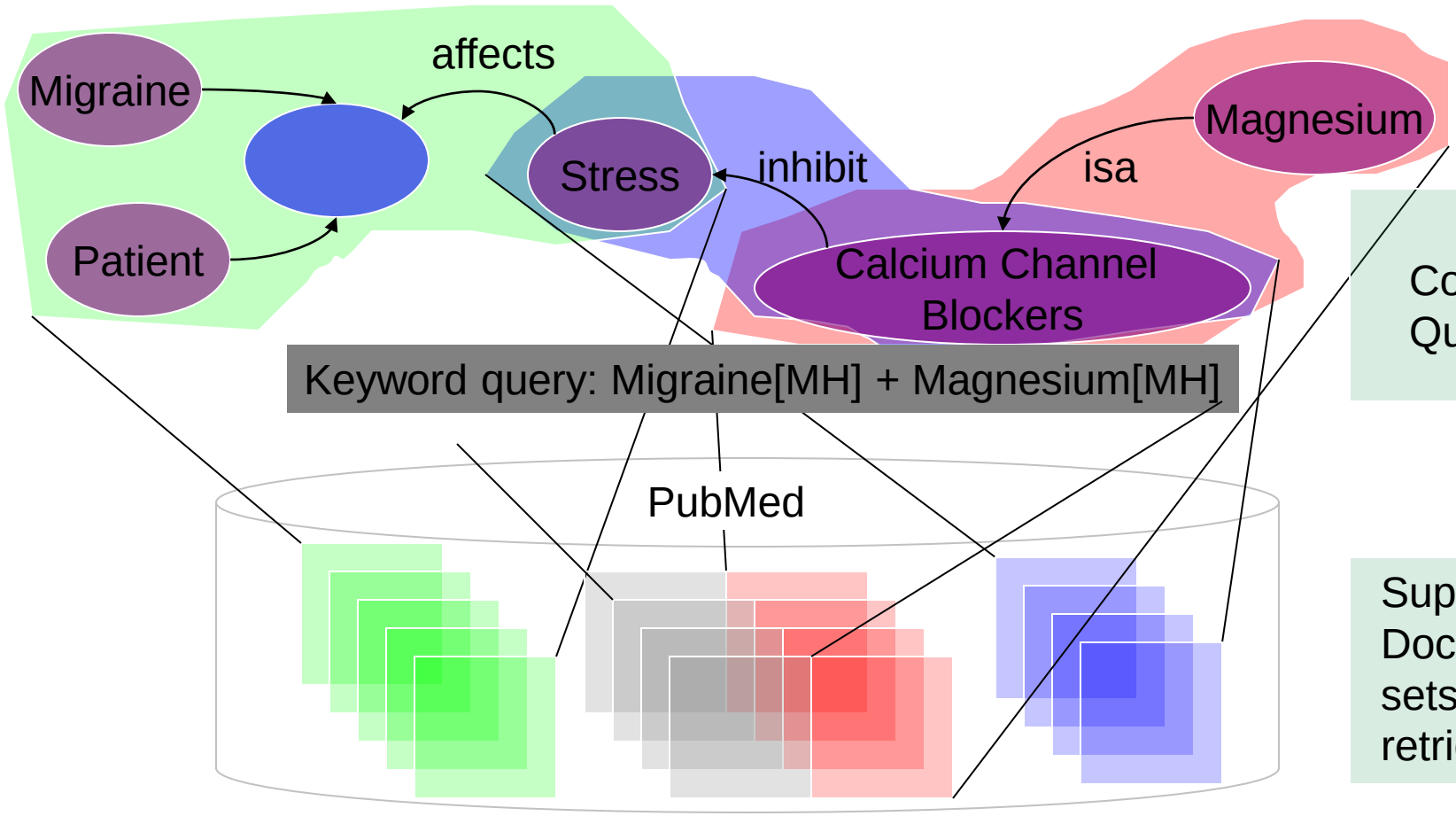
A second List of the Grand Masters of the Priory of Sion that included the names of Roger Patrice Pelat and Thomas Plantard appeared in 1989, but it should not be confused with the

Done

Auto Copy DISABLED Adblock

start Discovering Informati... KnowledgeDiscovery.... Drawing1 - Microsoft ... Yahoo! Messenger 5 Firefox 1:45 PM

Hypothesis Driven retrieval of Scientific Literature



Semantic Application in a Global Bank

Aim: Legislation (PATRIOT ACT) requires banks to identify 'who' they are doing business with

Problem

- Volume of internal and external data needed to be accessed
- Complex name matching and disambiguation criteria
- Requirement to 'risk score' certain attributes of this data

Approach

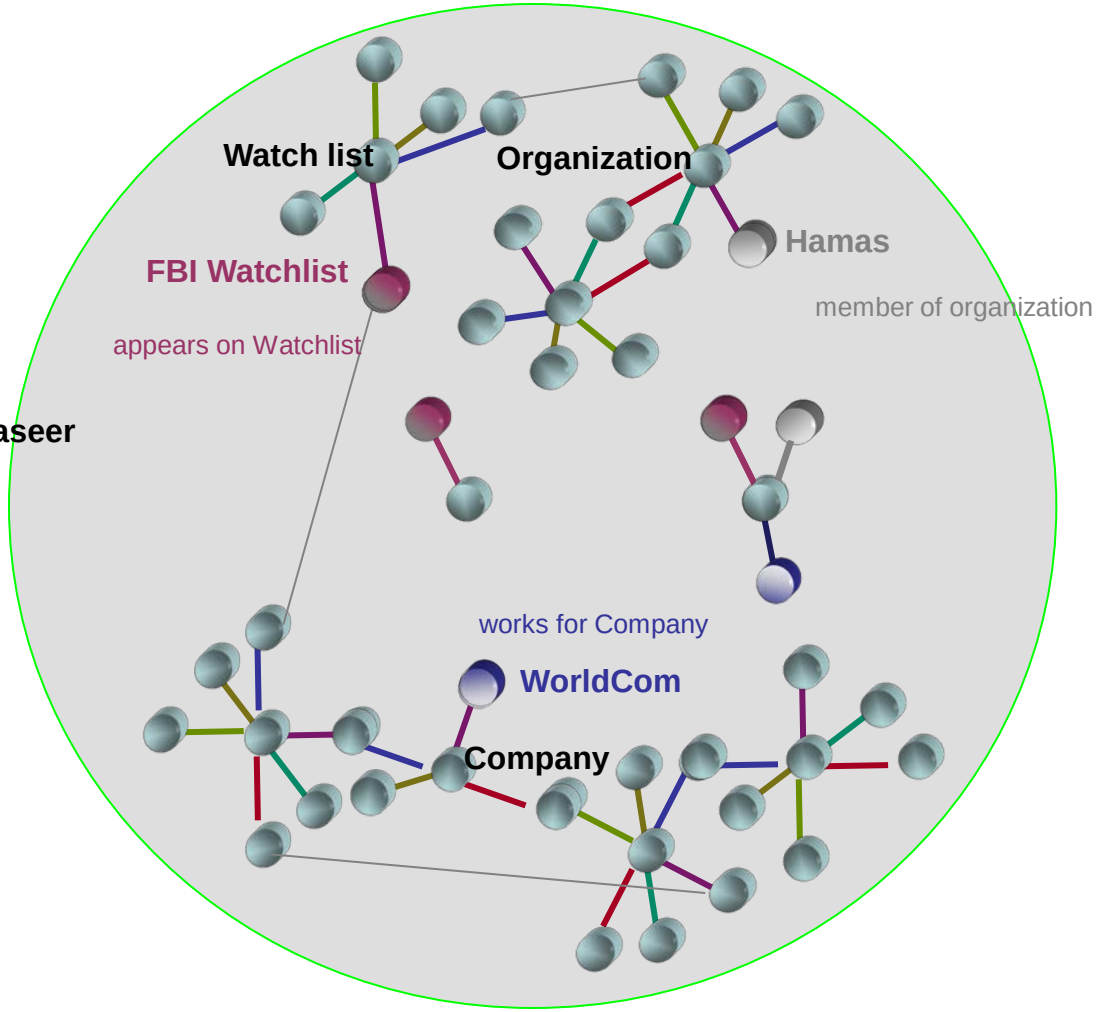
- Creation of a 'risk ontology' populated from trusted sources (OFAC etc); Sophisticated entity disambiguation
- Semantic querying, Rules specification & processing

Solution

- Rapid and accurate KYC checks
- Risk scoring of relationships allowing for prioritisation of results; Full visibility of sources and trustworthiness

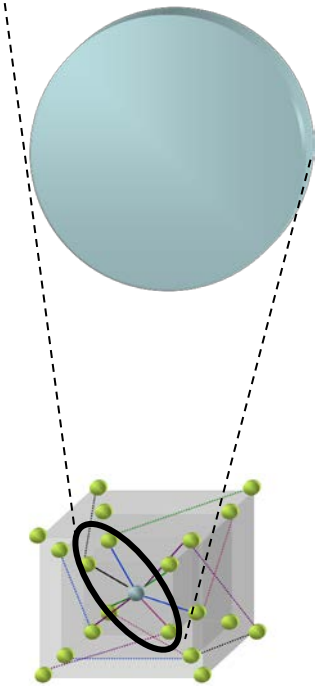
The Process

Ahmed Yaseer

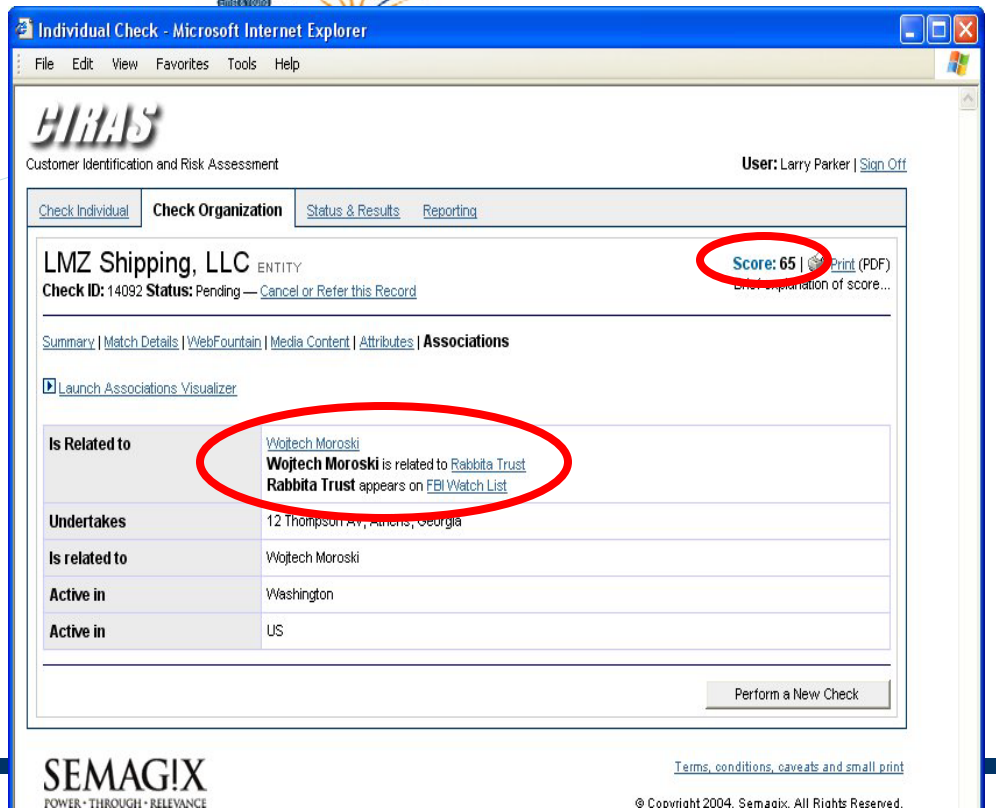


Ahmed Yaseer:

- ***Appears on Watchlist 'FBI'***
- ***Works for Company 'WorldCom'***
- ***Member of organization 'Hamas'***



Establishing New Account



Example of Fraud Prevention application used in financial services

Active Semantic Medical Records

(operational since January 2006)

Goals:

- Increase efficiency
- Reduce Errors, Improve Patient Satisfaction & Reporting
- Improve Profitability (better billing)

Technologies:

- Ontologies, semantic annotations & rules
- Service Oriented Architecture

Thanks -- Dr. Agrawal, Dr. Wingeth, and others

- Semantic Browser: contextual browsing of PubMed

- Passenger Threat Analysis
- Need to Know -> Demo
- Financial Irregularity *

* a classified application

Primary Funding by **ARDA**, Secondary Funding by NSF

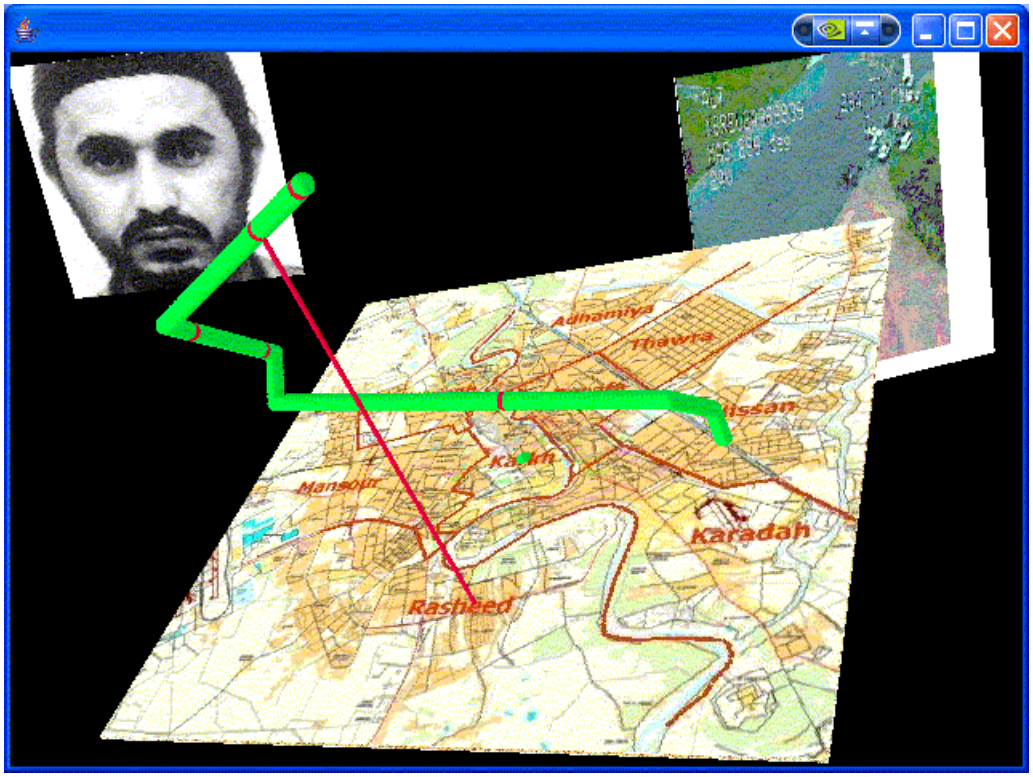
Aim

- Visualization with interactive search and analytics interface

Problem (examples)

- Need for intuitive visual display of semantic analytics showing "connections between the dots" between heterogeneous documents and multi-modal content
- Need for graphical tracking and association of activities to discover semantic associations between events using thematic and topological relations

Semantic Visualization: EventTracker



- Visualization of association unfolding over time
- Integration of associated multimedia content
- Separate Temporal, Geospatial, and Thematic ontologies describe data
- DEMO

- World class research center- coupled with daytaOhio for tech transfer and commercialization
- Core expertise in
  - data management: integration, mining, analytics, visualization
  - distributed computing: services/grid computing
  - Semantic Web
  - Bioinformatics, etc.
- With domain/application expertise in Government, Industry, Biomedicine
- Member of World Wide Web Consortium and extensive industry relationships

Invitation to work with the center

- Interns, future employees
- Targets research & prototyping
- dataOhio supported technology transfer/incubator
- Joint projects (e.g., C3 funding from DoD)
- Guidance on using standards and technologies

"increasing the value of data to you and your customers"

Introducing

http://knoesis.wright.edu